



Shedding light on causation between ISO 9001 and improved business performance

Shedding light
on causation

687

Gavin P.M. Dick

Kent Business School, University of Kent, Canterbury, UK

Iñaki Heras

The Basque Country University, San Sebastian, Spain, and

Martí Casadesús

University of Girona, Girona, Spain

Received 6 December 2006

Revised 22 November 2007

Accepted 19 March 2008

Abstract

Purpose – The adoption of the ISO 9001 Quality Management Systems Standard has proven to be a persistent and growing phenomenon in services and manufacturing, yet to date little research has been done that can indicate how far improved business performance can be attributed to it rather than counter-intuitive causes. The paper aims to examine the evidence for the causal links between quality management system certification and improved performance in the empirical literature.

Design/methodology/approach – A method is proposed for testing how far performance improvement can be attributed to quality management system certification and how far attribution to other causes applies. This method is illustrated on a longitudinal study and then utilised to interpret the findings of other longitudinal studies.

Findings – It is concluded that although there is some evidence to indicate that quality management system certification has some causal influence on business performance, there is also evidence for the existence of a substantial mechanism whereby better performing firms self-select to adopt certification. Possible causes for this mechanism are discussed.

Research limitations/implications – The existence of a self-select mechanism has profound implications for interpreting business performance achievements associated with quality management system certification because the benefits found may well be inflated by its presence. The authors suggest that richer theory is needed that can incorporate bi-directional influences and new research is needed to explore the underlying causes of adoption selection effects.

Originality/value – The paper provides researchers with a method for testing and discussing causation influences on results. It provides evidence that a substantial part of the association found in the research on quality management system certification and business benefits may be due to counterintuitive causes.

Keywords Performance management, ISO 9000 series, Cause and effect analysis, Quality

Paper type Research paper



Introduction

Although most “new” ideas in management have short life spans and are discarded when eclipsed by the next fad (Carson *et al.*, 2000), adoption of third party accredited ISO quality management system standards has proven to be a persistent and growing phenomenon. Its persistence suggests that it is not simply another management fad but will remain an influential global management meta-standard (Uzumeri, 1997).

International Journal of Operations &
Production Management
Vol. 28 No. 7, 2008
pp. 687-708
© Emerald Group Publishing Limited
0144-3577
DOI 10.1108/01443570810881811

Despite the high cost of achieving and maintaining registration to the ISO 9001 quality management system standard, more than 750,000 organizations in 161 economies have made the investment (ISO, 2005). ISO 9001 Registrars make bold claims for the business benefits of quality management system certification, for instance in the USA, ANAB (2005) claim 16 benefits from quality management system certification including increased operational efficiency, cost savings from less rework, customer satisfaction, competitive edge, perceived higher quality and increased market share. In Europe similar claims are made (Breeze, 2004). This raises the question of whether the claims for attribution of improved performance to quality management system certification are valid. This motivates us to explore whether there is evidence for benefits and in particular what evidence there is to prove whether mechanisms other than certification to the ISO 9001 standard could be the cause.

In this paper, we propose a method for identifying causal direction of performance and contrast its results with those for the cross-sectional analysis methods usually employed. We demonstrate that the proposed method can lead to very different results and conclusions to those obtained from cross-sectional analysis methods. We find that reverse attribution (better performance preceding quality management system certification) is a major mechanism that explains the superior performance of certified firms found in our earlier study (Heras *et al.*, 2002a). The method is then used to interpret the results of previous empirical studies. These analyses cast doubt on the causal inferences being drawn from the literature that finds an association of ISO 9001 accreditation with better business performance.

The paper looks first at the causal links between quality management system certification and improved performance before looking at evidence for them in the empirical literature. We then discuss methods for attributing causal direction and demonstrate their use on previously published data. We conclude with a discussion of the implications of our analysis for the attribution of performance in studies examining quality management system certification to ISO 9001 (hereafter referred to as QCert).

Literature

A quality and business causal model

Although there is general agreement in the literature on the association between quality and performance, we need to note that there is little commonality in how the literature measures business performance or defines quality (Sousa and Voss, 2002).

The literature is in broad agreement on the potential causal chain between improved quality systems and better performance. Both Garvin's (1984) quality model and Deming's (1986) imply that, as quality improves, waste is eliminated, costs are reduced, and financial performance improves. In the context of the ISO 9001 quality management systems standard the causal links can be extended as follows. A certified quality management system can achieve an increased emphasis on quality (Dick *et al.*, 2000) leading to less waste and duplication of effort, and improvement in product quality. This means there are lower costs and less customer attrition which leads to increased sales volume, while lowering the average cost of acquiring new business. These in turn lead to improved profitability from a combination of lower cost of production, lower sales expenses and scale economies from greater sales volume. Indeed, even if not all the quality benefits materialize, the possession of the "Quality Badge" alone could lead to increased sales opportunities and so, improve profitability

from increased sales volume. This causal model of improvements flowing from quality management system certification to improved business performance is shown in Figure 1. We next examine the evidence in the literature for the performance benefits shown in Figure 1 and whether they are caused by quality management system certification.

ISO 9001 quality management and performance literature

Here, we review the empirical work in peer-reviewed journals from 1990 to 2006 that include reference to ISO 9000 or ISO 9001, and performance or benefits. A six-stage approach to selection of articles was used. Initial screening of the 2000 or so electronic search listings excluded materials that were not in peer-reviewed journals, followed by a relevance screening to exclude articles that did not explicitly measure business benefits or performance variables. At this stage it was found that there were many studies reporting expectations of increased market share and improved product quality from ISO 9001 implementation (Ebrahimpour *et al.*, 1997), but there were less than 100 empirical studies on the business performance benefits actually achieved.

Of these articles only those that explicitly measured benefits that could be related to the model of Figure 1 were included. These were measurement of any of the following: waste, costs, better quality, higher sales, market-share or profitability ratios. A few articles were included that combined lower waste and lower cost as these can be viewed as closely related operational measures. Next, the research methods of each article were assessed and only studies that reported the statistical significance of the results and had sample sizes with sufficient statistical power were chosen. To determine statistical power we used Cohen's (1988, p. 31) table that calculates that samples larger than 76 are needed to ensure that a relationship >0.4 will be detected at a significance level of 0.05 and a power of 0.8. This screening ensured that we avoided studies whose sample sizes are unlikely to detect statistically significant effects that are medium in magnitude and very unlikely to detect small effects.

Next, papers that included firms registered after 2000 were excluded, so that the findings could be viewed as being uninfluenced by the major quality standards revisions (ISO 9001:2000) that were applied from 2001 onwards. Finally, where an author had generated multiple publications from the same research data the paper that best matched our criteria was selected. Clearly, this process cannot claim to have

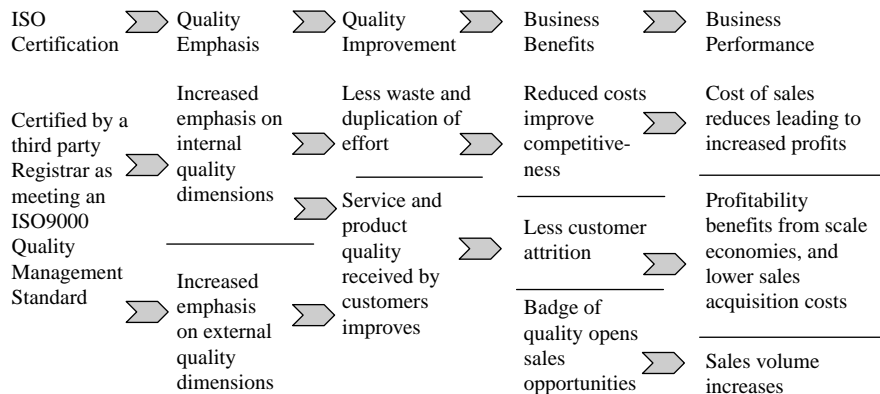


Figure 1.
The expected links
between management
system certification to ISO
9000 standards and
business performance

captured every item of relevant research but it can be viewed as a substantial sample of the literature, which is unlikely to have any systematic bias in its selection.

This methodology resulted in a set of 26 research study results which may seem a small number compared to the apparently vast research output relating to ISO 9001 but it is comparable to findings of Ahire *et al.* (1995) who found only 29 empirical articles from the 226 on total quality management and business performance that they reviewed. Our analysis breaks the quality management system certification (QCert) papers we have selected into two groups. The first of these consists of “snapshot” (cross-sectional) studies or studies analysed on a cross-sectional basis that provide evidence of statistical validity. Peer-reviewed papers of this type started to appear in 1997 and continue with many of these using intervening variables on the firms’ motivation for pursuing QCert to explain when performance gains were achieved. We start by briefly summarizing the findings of this first group in Table I before moving on to analyse in greater detail the second group that will cover longitudinal studies that have the potential to indicate causal direction.

Table I provides a summary of the first groups’ findings and shows the study’s first named author and whether their results support or not an association between QCert and the listed benefits. We start by summarising the studies in the upper part of Table I that do not include the influence of intervening variables. Overall, the upper part of the table indicates that the strongest findings are for the internal benefits of less waste (five from seven studies) and lower cost (five from nine studies). Support for external benefits is less compelling with better quality found in four from eight studies while only four from 11 studies found higher sales/market share and four from ten higher profits. In the lower part of the Table I, we see the research that uses “motivation” intervening variables where the common theme is whether the motives for adoption are internal or developmental rather than external motives (such as marketing benefits or customer pressure). The lower part of Table I shows that when “internal/developmental” motives variables are present then most studies report a range of benefits from QCert. All eight studies find internal benefits (lower waste and/or lower costs) while four from six studies report better quality. Results for higher sales/market share (three from six studies) or profitability (one from two studies) are mixed as only half show benefits. However, what is striking is that when internal/developmental motives for QCert adoption are absent, all the eight studies find that there are no business benefits to report.

Overall, the evaluation of the links between QCert and improved performance reveals that there is evidence in the field’s empirical research to suggest that the broad range of benefits shown in Figure 1 are possible but uncertain unless motivation for pursuing QCert is for internal or developmental reasons. However, caution is needed in inferring that certification is the cause of the benefits found since the methodologies that are used in all the studies we have reviewed so far assume a positive direction of causation, i.e. that quality system certification leads to improved performance. Could these studies’ assumption of only forward causality between ISO 9001 certification and improved business performance be erroneous? Could it be that reverse attribution also exists, i.e. that better business performance precedes QCert and is being mistakenly attributed to QCert? In other words, could it be that organizations with above average business performance tend to pursue QCert more than less profitable firms? This would explain or inflate the better performance found in the presence of QCert.

First named author ^a	Internal benefits		External benefits		Higher sales or market share		Profit higher	
	Lower waste	Lower cost	Better quality					
Buttle (1997)	Yes	Yes			Yes		Yes	
Terziovski (1997)	No	No	No		No	No	Yes	
Simmons (1999)						No	Yes	
Huarng (1999)	Yes	Yes	Yes		Yes		Yes	
Sun (2000)	Yes	Yes	Yes					
Prabhu (2000)	Yes			No				
Lima (2000)					Yes			No
Singles (2001)			No	No		No		No
Santos (2001)			No	Yes		No		No
Tsekouras (2002)								No
Chou-Chua (2003)	No	No	No		No	No		No
Terziovski (2003)			Yes		Yes			
Dimara (2004)						No		No
Briscoe (2005)	Yes	Yes				No		
Tzelepis (2006)		Yes					Yes	
Totals	5	2	5	4	4	4	7	4
	Yes	No	Yes	No	Yes	No	Yes	No
<i>Intervening variable^b</i>								
Jones (1997)	Developmental	Yes	Yes		Yes	Yes		
Brown (1998)	Internal motivation	Yes		Yes	Yes			
Abraham (2000)	Leadership, etc.	Yes	Yes		Yes			
Yahya (2001)	Developmental	Yes		No	No	No		
Singles (2001)	Internal motivation	Yes	Yes	Yes	Yes	Yes		
Yeung (2003)	Internal motivation	Yes	Yes	No	No	No		
Terziovski (2003)	Internal motivation	Yes						
Naveh (2005)	Catalyst for change	Yes	Yes			No	No	
Totals	If variable present	5	0	5	0	4	2	3
	If variable absent	0	5	0	5	0	6	0
		Yes	No	Yes	No	Yes	No	Yes

Notes: "Yes" indicates supporting results reported; "No" indicates no support in results; ^afull citations are given in the reference section; ^bvariables measuring whether accreditation was for internal motives or externally driven

Table I.
Summary of research on
ISO 9000 certification and
business benefits

To examine this causation question we now examine in detail the second group of studies that covers four research articles that used research designs that could provide evidence of causality. Each of these longitudinal studies starts at the point when registration to the ISO 9001 standard began to expand in the country or sectors examined.

The first was Häversjö's (2000) longitudinal analysis of the returns on capital employed of Danish companies between 1989 and 1995. In this study, the 871 companies were compared with a control group of 644 firms matched by size, to see if the abnormal rate of return on capital employed improved after registrations. Häversjö's longitudinal results (Häversjö's, Table I, p. 48; summarized here in Figure 2) showed that the average financial performance of the certified organizations was superior to the non-certified organizations both before and after their registration but

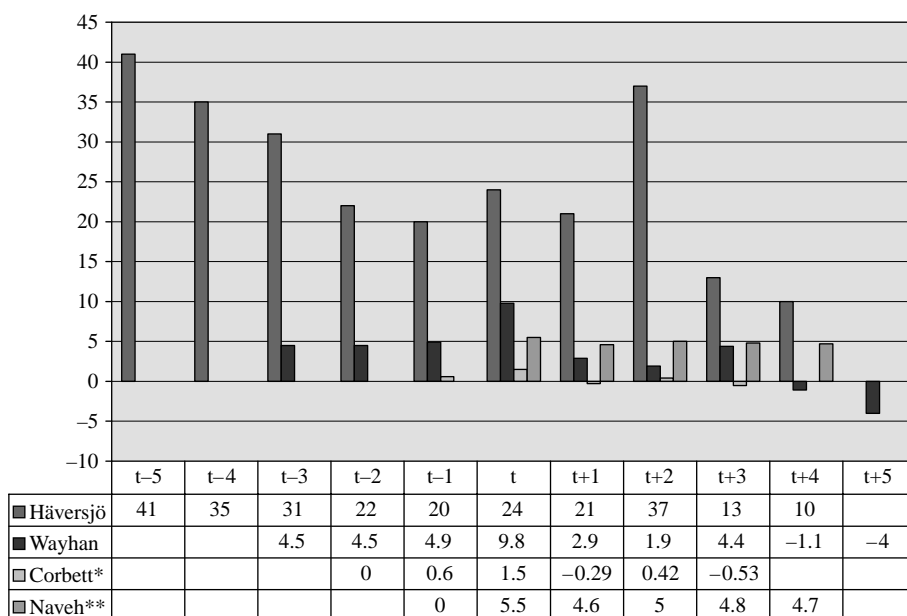


Figure 2.
A summary of the
abnormal profitably
findings in longitudinal
studies of QCert firms

Notes: t is the year of quality certification, $t + 1$ is one year after quality certification, etc; *Corbett ROA standardized on $t - 2$ year; **Naveh ROA standardized on $t - 1$ year

that no consistent statistically significant pattern of post-registration performance gains could be detected.

The second article that used a research design that could provide evidence of causality is Wayhan *et al.*'s (2002) analysis of the performance of 96 organizations in the USA between 1990 and 1998. Their table of results (Wayhan *et al.*'s Table I, p. 225; summarized here in Figure 2) also shows that the 48 registered organizations had a consistently better return on assets employed, both before and after registration, compared to a control group of 48 non-registered organizations who were matched by industry and size to the certified companies. As with Häversjö no significant post-registration performance gains can be detected.

The third examined the performance of 544 US firms between 1987 and 1997 (Corbett *et al.*, 2005). In their study, to avoid the influence of reverse attribution control groups were constructed following the methods of Barber and Lyon (1996). These involved individual matching and portfolio matching to, "industry-firm size" "industry-firm size-ROA", "industry-ROA", to achieve six control sets. Their findings indicate that differences in abnormal performance in ROA tend to be highest when compared to a control of one to one matching by industry and ROA (Table IV, p. 1052); but all the differences regardless of the control used show achievement of modest year-by-year gains (0.5 to 1.5 per cent) post QCert, few of which reach statistical significance. However, when these gains were aggregated, which increases the power of the tests, all of the control group sets showed that these small gains became statistically significant.

The fourth is Naveh and Marcus's (2005) examination of 313 US firms between 1990 and 2000. Like Corbett *et al.* (2005) they used Barber and Lyons (1996) method to eliminate the influence of reverse attribution. They paired individual firms by, "industry", "industry-firm size", "industry-firm size-ROA" and "industry-firm size-ROA-Stock price performance". Regardless of control group type they found only non-significant yearly ROA gains for all of the five years following registration. However, like Corbett *et al.* they found that when these gains (which were more substantial than those found by Corbett *et al.*) were aggregated over the post registration five years these gains became statistically significant. This applied to all of the control groups.

Before drawing conclusions on these studies we need to consider whether different control group selection methods affect the reliability of results. To provide insights into this question we need studies that provide results for multiple non-certified control groups each of which is constructed by alternative methods. Fortunately two of the longitudinal studies that we have reviewed do this. Naveh and Marcus' (2005) findings indicate that industry matching is the most reliable criterion since they found little difference in the abnormal ROA of certified firms over a 60-month period when matching was done with an "industry matched control group" (ROA 5.23 per cent) or an "industry and size matched control group" (ROA 5.35 per cent). Similar conclusions are found in the accounting literature where matching by industry is considered the dominant criterion because size in addition to industry provides little or no benefit (Barber and Lyon, 1996). However, caution is indicated from the results of Corbett *et al.* (2005) whose abnormal ROA over a five-year period shows that controlling for size (assets) in addition to industry does produce different results (ROA 8.98 per cent) than industry matching alone (ROA 5.82 per cent). From this it can be concluded that including a size control in addition to matching by industry is a prudent approach. This brings us to the question of how close should industry be matched in the control group?

Wayhan *et al.*'s method appears strong since it uses a four-digit SIC industry criterion for their control group that provides a close industry match along with size matching by assets. Naveh and Marcus (2005) uses a three-digit SIC criterion where possible while Corbett *et al.* (2005) used a looser two-digit SIC criterion as they wished to avoid discrepancies at the three- and four-digit level that have been found in their database by previous researchers (Guenther and Rosman, 1994). Although it is clear that each of the longitudinal studies uses different detailed selection methods for their control groups all apart from Häversjö have used a systematic approach that attempts to align their control group as closely as possible to the industries of their QCert sample along with a control for the firms' size. Thus, each study may have a distortion in their results because of methodological differences that in absolute terms makes them not directly comparable. However, what is noticeable in both Naveh and Marcus, and Corbett *et al.* studies are that regardless of the control group used, the year by year trend in the detailed results remains reasonably consistent.

The next question is whether it is better to use one-to-one matching in the construction of the non-certified control group rather than using a portfolio approach where the certified firm is compared with a pool of similar firms. Corbett *et al.*'s study provides information on this as they contrast results for a "portfolio" and "one-to-one matched" method for each of their control group construction methods. Their detailed

results show that although there are some differences in results between portfolio and one-to-one matching, the trend in results within the groups is similar.

Therefore, regardless of differences in control group construction it appears that the trend in results for each control group choice remains reasonably consistent year by year. This is because these four studies use repeated measures so any distortions due to the method of selection of the control group are consistent, therefore allowing reliable interpretation of inter-year performance differences for each study. So we must conclude that comparing the absolute level of abnormal ROA between one study and another can be an unreliable measure because their comparator will be affected by differences in control group construction method. However, the year by year differences within each study can be viewed as reliable indicators because they are not affected significantly by the choice of control group construction method. Thus, we can reliably analyse the influence of QCert on performance over time for these studies.

The results of these four studies are shown in Figure 2 (here we report Corbett *et al.*'s and Naveh and Marcus's findings against their "industry-firm size-ROA" matched control groups since their chosen reporting method aims to standardize ROA to show abnormal ROA returns above those achieved pre-certification). In Figure 2, it can be clearly seen that there is no discernable improvement trend in profitability post-certification in Häversjö or Wayhan *et al.*'s studies while Corbett *et al.*'s post-certification performance indicates only small changes compared to the steady and worthwhile performance differential seen in Naveh and Marcus's firms.

From Figure 2, it is clear that the pre-certification profitability of Häversjö's and Wayhan *et al.*'s certified firms is better in all the years leading up to certification compared to their non-certified control groups. This better than average pre-certification profitability is also noted by Corbett *et al.* (2005, pp. 1051, 1057) and Naveh and Marcus (2005, p. 19. Table VIa) and eliminating its effects on their results is one of the prime objectives of their methodologies. So, taken together these longitudinal studies do provide evidence that adopters of QCert tend to be firms with above average performance prior to their certification. In other words better performing firms have a tendency to self-select to adopt certification. However, of the four longitudinal studies that could indicate that QCert leads to improved profitability only two do so. Naveh and Marcus provide evidence of profitability benefits of a meaningful effect size while Corbett *et al.* find statistically significant gains, but on a more modest scale.

This tendency towards superior performance prior to QCert is a cause for concern since it suggests that the cross-sectional studies/analyses, reviewed earlier, that report performance benefits from QCert may be attributing to QCert benefits that are in part or primarily due to the better performance preceding certification.

Clearly, attributing performance to QCert (or any other management initiative) is more complex than might first appear, and our review of the empirical literature suggests that there is a paucity of research designs that can show that performance benefits found can be safely attributed to QCert and none that explicitly set out to separate the influence of reverse attribution from that of quality certification. With this in mind we next discuss a method for testing attribution of performance, demonstrate its use and then relate our findings to those in our literature review.

Methodology

Attribution testing

Ideas on causation have exercised philosophers since Aristotle but perhaps the most appropriate modern regularity theory for use in the management field of enquiry is that of a cause being a sufficient condition for the occurrence of some effect with the rider that the cause must precede the effect and other possible explanations are eliminated (White, 1990). In practice, in the social sciences, causality is usually accepted in empirical research as requiring three conditions:

- (1) that there is an association found between variables that logically might influence one another;
- (2) the causal variable must produce its influence before the outcome occurs; and
- (3) other possible explanations must be eliminated such as a third variable that influences both variables (Blaikie, 2003).

Here, we need to be careful not to claim too much when we use the word causation. We acknowledge it is not possible to test for true causation through statistical methods; instead our focus is on differentiating the forward and reverse influences that can be attributed as causes of performance improvement, hence our use of the term attribution testing. In other words, we wish to separate performance effects that are due to counter-intuitive causes (that we will label as “reverse attribution”) from effects that can be attributed to the dependent variables(s) being studied.

QCert has been shown in our literature review to have a chain of influences that might be a sufficient condition for the occurrence of better financial and sales performance. In other words, we have a plausible sequence of casual relationships (Figure 1) that we can view as mechanisms that can explain why QCert could cause improved financial performance and we have found associations between them that indicate that a relationship exists between them. However, for causation to be attributed we also need to satisfy the other two conditions. We need to show that better performance did not precede QCert and that there are no other explanations for it.

So how can this be operationalised? To do this we need research designs that go beyond the dichotomous idea of comparing certified firms with those that are not-certified, by splitting from the non-certified firms those firms that will be certified in the future. These “not-yet-certified” firms are firms that are not certified in the year being analysed but become certified firms later in the longitudinal study. Thus, in our design we have three cross-sections, not-yet-certified, certified and a control group of non-certified firms. Figure 3 shows these three groups and their relationships.

To test whether performance improvement is entirely due to QCert we would need to show that three conditions are satisfied. Firstly, we would need to show that the effect exists in the presence of QCert but is absent when QCert is not present. Secondly, we would need to show that the effect does not precede QCert, and thirdly we would need to demonstrate that the effect magnitude of the QCert influence on performance was substantial. So we would need to test for the effect by showing that certified firms [x] had better performance than non-certified firms [y]. We would need to establish that firms prior to their certifications that we label “not-yet-certified firms” [z] had similar performance to non-certified firms. Finally, we would need to test for the magnitude of the influence (effect size) by showing that certified firms had better

performance compared to not-yet-certified ones. Thus, full causal attribution to QCert requires: $x > y$, $z \equiv y$, $x > z$.

However, if we find that not-yet-certified firms and certified firms have better performance than non-certified firms and there is little or no difference in performance between not-yet-certified firms and certified firms then a full reverse attribution mechanism is found. This indicates causes other than quality QCert are responsible for the better performance. Thus, full attribution of better performance to other causes requires: $x > y$, $z > y$, $x \equiv z$.

If better performance is found in both not-yet-certified and certified than non-certified firms, and certified firms have better performance than not-yet-certified firms, then we have co-causality of better performance. Some of the performance differences can be attributed to QCert with the rest being attributed to reverse attribution. Thus, the performance attributable to QCert = $(x - y) - (z - y)$ and that for reverse attribution = $(x - y) - (x - z)$.

Therefore, we used three significance tests to determine the relationships between performance improvement and QCert:

- (1) T1, this is the size of the combined effects of causal attribution to certification and reverse attribution [aggregate effect size test $(x - y) > 0$].
- (2) T2, this is the size of the effect of reverse attribution [reverse attribution effect size test $(z - y) > 0$].
- (3) T3, this is the size of the causal attribution to certification [certification attribution effect size test $(x - z) > 0$].

Figure 3 graphically shows these tests. Using the causality logic, we can see that T1 tests for a significant effect associated with the cause(s), while T2 is a causal attribution test which determines the magnitude of reverse attribution. Finally, T3 is a test of the magnitude of the effect attributable to the independent variable QCert.

At this point we need to make it clear that we are not suggesting that these tests are the ultimate solution in attribution testing, for that more complex and costly research designs are required (Pearl, 2005). Rather, the simple tests proposed here represent the minimum needed to test for the potential influence of reverse attribution in the

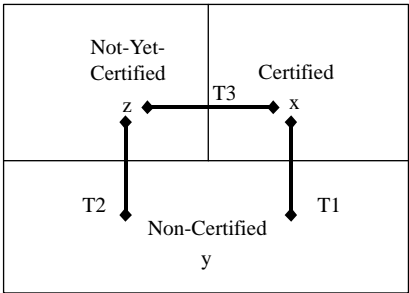


Figure 3.
Tests for performance
attribution

Notes: T1 tests for overall effect, T2 for reverse attribution, T3 for attribution to certification; attribution to QCert = $(x - z) > 0$; reverse attribution = $(z - y) > 0$

interpretation of performance attribution. At this point we need to emphasise that the primary purpose of this paper is not to analyze the pros and cons of different methods that can be used to “safely” attribute performance to an independent variable (which we acknowledge that the Barber and Lyon methods used by Corbett *et al.* and Naveh and Marcus that were reviewed earlier do achieve) but to explore the possible influence of reverse attribution in the interpretation of the many studies whose results can only show association. However, the testing method we propose does have some advantages over the prior performance matching methods of Barber and Lyon (1996).

Foremost is that the attribution testing method reports the effects attributable to the independent variable and any reverse attribution. Knowing the effect size of reverse attribution is important in its own right if we are to present findings in a broader theoretical frame where the influence of unknown effects is acknowledged and propositions derived that may explain their possible origins. In addition the method proposed is versatile and can be used on small data sets, unlike the panel-data matching methods (Barber and Lyon, 1996) used by Corbett *et al.* and Naveh and Marcus that require very large longitudinal data sets. Finally, the attribution testing method we propose can be appropriate in situations where data is only available at limited intervals where the Barber and Lyon approaches would not work.

In summary, the method proposed avoids some of the demands of panel matching methods but does introduce a drawback associated with any research that uses a cross sectional control group. This enables the influence of firm size and industry selection effects to be controlled for but there remains the risk that the control group can be unrepresentative in terms of its profitability. However, the method’s strength is that it provides the researcher with a metric for the effect size of reverse attribution.

A test of the method

The research data that we use to demonstrate the attribution testing method comes from the Basque Autonomous Community, which is considered, with Madrid and Cataluña, to be one of the regions in Spain where ISO 9001 registrations are concentrated. The dataset is identical to that used by Heras *et al.* (2002a). In that paper, however, it was analyzed using a T1 type test that can only indicate association.

The data for that study were gathered from the Ardán database that is one of the most comprehensive in Spain for economic and financial information. The data are recorded from, among other sources, the profit and loss accounts and balance sheets that companies must submit to the Mercantile Register.

For the analysis, two samples were drawn from the database. The first was a sample of 400 ISO 9001 certified companies (in the manufacturing, construction, retail and services sectors), the first of which were registered in 1995. The second was a sample of 400 non-certified companies matched by industrial sector to four digits of the CNAE code (the Spanish equivalent of SIC codes). Data were available for the years 1994, 1995, 1996, 1997 and 1998, and included the sales revenue for each accounting year, as well as the profitability ratio (ROA, the ratio of net profit before interest and tax on total assets). In addition, for the certified companies, the data set included information on their last quality certification registration date. This information on registration dates was checked with the registration bodies and, where necessary, with the companies to ensure that the date used was the true date of the firm’s initial registration to ISO 9001.

Possible sources of bias in the two samples were checked. We noted that certified firms had on average larger sales turnover than non-certified firms did. To test that any difference in profitability of the certified companies was not a direct result of their larger sales, we used the *z*-test of proportions, with a level of significance set at 0.05, as well as a *t*-test for differences in means. Both these calculations indicated that there was no statistically significant effect of turnover on ROA.

Likewise, to see if industry selection effects existed for ISO 9001, the average profitability ratio for all the sectors (manufacturing, construction, retail and services) for all years was calculated to establish whether any sector differences between the samples and the control were creating a bias in the results. No statistically significant differences were identified using *t*-tests. Therefore, we may be confident that any differences between ISO certified and non-certified companies found are not related to the firms' size or sector distribution of the two samples.

In the study we used the registration year to split the not-yet-certified from the certified companies since we found no evidence of any increase in firms' performance in the one or two years prior to certification in the Figure 2 data from the Häversjö or Wayhan *et al.* studies or in our earlier work that used an event study method on our data (Heras *et al.*, 2002b).

In summary, the research design consists of three samples of firms: certified, not-yet-certified and non-certified for each of the five years, and two variables, sales growth, and return on total assets employed (ROA).

Findings

We start by briefly presenting the findings of the original longitudinal study (Heras *et al.*, 2002a). These provide a starting point that allows later comparison with the attribution testing method results. In the original paper a dichotomous split was made between certified and non-certified firms (not-yet-certified firms being excluded from the analysis). The results for the two samples average year on year sales growth are presented in Table II. The findings indicate that certified firms achieved substantially greater cumulative average sales growth (56 per cent) than non-certified firms (40 per cent) during the five years, with two out of the four years being statistically significant.

A similar picture emerges for profitability (Table III) with certified firms enjoying better average profitability (ROA) than non-certified firms over the five-year period with their average ROA being 8.67 per cent compared to non-certified firms' 6.89 per cent. Here, two out of the four years show statistically significant differences.

These sales and profitability results provide good evidence for sustainable improved performance being associated with QCert (given that the tests for company size bias and industrial sector ISO 9001 selection effects showed these had no influence). However, as often seems to be the case in this field, practitioners' intuition is

Table II.
Average sales growth
in ISO 9000 certified and
non-certified companies

	1994-1995	1995-1996	1996-1997	1997-1998	Cumulative
Non-certified (per cent)	13.88	5.30	11.77	8.70	40
Certified (per cent)	*25.69	*10.40	10.84	9.31	*56

Note: *Indicates that the Mann-Whitney test is significant (*p* < 0.05)

to claim causality rather than association. For instance, the results shown here in Tables II and III have often been quoted by the CEO of the British Standards Institute (the lead registrar for ISO 9001 in Great Britain) as evidence for Certification achieving significant sales and profitability benefits (Breeze, 2004). If causality is claimed on the basis of a T1 effect test it is being assumed that not-yet-certified firms have similar performance to non-certified firms (T2: $z \equiv y$) which in turn implies that a T3 test ($x > z$) would show similar gains to a T1 test ($x > y$).

To see if these implications are valid we now examine the same data set but include in our findings the results for not-yet-certified firms (these are firms that will be certified later in our longitudinal analysis). The findings for sales growth are shown in Table IV. Shown alongside the percentage sales growth for each of the years are the significance level results for Mann-Whitney U tests for the attribution tests we described earlier. Overall sales growth is significantly better for certified (56 per cent) and not-yet-certified (60 per cent) than non-certified firms (40 per cent) (so T1 result $x > y$ and T2 result $z > y$) whilst for T3 QCert attributable plus and negative differences are found in individual years with a cumulative -4 per cent sales growth. None of these T3 results are statistical significant so the result is $x \equiv y$. Therefore, the attribution tests meet the conditions for reverse attribution, i.e. that firms had greater sales growth than their peers before QCert but achieve no additional significant sales after it.

The findings for profitability are shown in Table V. Overall ROA is significantly better for certified (8.67 per cent) and not-yet-certified (8.96 per cent) than non-certified firms (6.89 per cent) so T1 result is $x > y$ and T2 results is $z > y$. The ROA cumulative

	1994	1995	1996	1997	1998	Average
Non-certified	5.50	7.08	5.70	6.76	7.78	6.89
Certified		6.37	*8.48	8.29	*9.66	*8.67

Notes: Figures are the per cent return on total assets employed; * indicates that the Mann-Whitney test is significant ($p < 0.05$)

Table III.
Average profitability
(ROA) for the certified
and non-certified
companies

	1994-1995	1995-1996	1996-1997	1997-1998	Cumulative
Non-certified (per cent)	13.88	5.30	11.77	8.70	40
[SD] n	[26] 400	[20] 399	[22] 398	[22] 399	
T1 certified (per cent)	*25.69	*10.40	10.84	9.31	*56
[SD] n	[23] 50	[21] 135	[20] 239	[22] 235	
T2 not-yet-certified (per cent)	*21.28	9.11	15.52	14.05	*60
[SD] n	[24] 328	[23] 308	[21] 190	[21] 49	
T3 QCert attribution (per cent)	ns 4.40	ns 1.29	ns (4.68)	ns (4.74)	ns (4)

Notes: [SD] is the standard deviation of the mean. n is the size of the sample. For T1 and T2, * indicates that the Mann-Whitney tests for certified or not-yet-certified firms' sales compared to non-certified firm's sales are significant ($p < 0.05$). T3 QCert attribution contrasts firms who are not-yet-certified (T2) with firms that are certified (T1). ns – indicates the T3 Mann-Whitney test indicates that any difference is not significant ($p > 0.05$)

Table IV.
Average sales growth
for the certified,
not-yet-certified and
non-certified companies

	1994	1995	1996	1997	1998	Average 95-8
Non-certified (per cent)	5.50	7.08	5.70	6.76	7.78	6.89
[SD] <i>n</i>	[9] 400	[10] 400	[10] 399	[10] 398	[11] 399	
T1 certified (per cent)		6.37	*8.48	8.29	*9.66	*8.67
[SD] <i>n</i>	0	[10] 50	[8] 135	[9] 239	[12] 235	
T2 not-yet-certified (per cent)	*6.57	*9.34	*8.44	*8.97	9.61	*8.96
[SD] <i>n</i>	[9] 354	[9] 328	[9] 308	[8] 190	[13] 49	
T3 QCert attribution (per cent)		*(2.97)	ns 0.04	ns (0.68)	ns 0.05	ns (0.29)

Table V.
Average profitability
(ROA) for the certified,
not-yet-certified, and
non-certified companies

Notes: [SD] is the standard deviation of the mean; *n* is the size of the sample. For T1 and T2, * indicates that the Mann-Whitney tests for certified or not-yet-certified firms' ROA compared to non-certified firm's ROA are significant ($p < 0.05$). T3 QCert attribution contrasts firms who are not-yet-certified (T2) with firms that are certified (T1). ns – indicates the T3 Mann-Whitney test indicates that any difference is not significant ($p > 0.05$)

differences attributable to QCert are small (–0.29 per cent) and not statistically significant so the T3 result: $x \equiv y$. Thus, as with the sales analysis, the overall ROA results meet the attribution test conditions for reverse attribution, e.g. that firms had greater ROA than their peers before QCert but show no additional profitability gains from it (given that the tests for company size bias and industrial sector ISO 9001 selection effects showed that these were not an influence).

If we compare these results with those in Tables II and III, we see a very different interpretation of the better results of QCert firms from those of the British Standards Institute (Breeze, 2004).

We can see that the causal attribution being claimed (Breeze, 2004) from T1 effect tests is incorrect since the practitioner's assumption of not-yet-certified firms being similar to non-certified firms ($z \equiv y$) that underlies their intuitive attribution of better performance is proved in our study to be false. Our results indicate that while for profitability, statistically significant ($z > y$) differences exist in all four T2 tests of performance preceding certification no significant effect magnitude following on from QCert ($x > y$) can be detected in any year in the T3 tests. It is now clear that any profitability gains found in the original research cannot be attributed to QCert but rather to a reverse attribution mechanism. A similar picture emerges when we compare sales growth in Table IV as no statistically significant gains are found in the T3 QCert attribution row for any year.

To check on the robustness of these findings they have been replicated using *t*-tests, all of which confirm our findings. In addition the test has been replicated using medians instead of mean values but no differences emerge that change the overall findings.

Discussion

Our findings show that when we tested our data using cross-sectional analysis methods such as those found in the majority of the empirical literature on ISO 9001, we also found a significantly better sales growth in certified companies than in the control group of non-certified ones. However, using our attribution testing method on the same data we found that none of these gains can be attributed to QCert. The tests indicate that there is no evidence to support any causal link between ISO 9001 registration and improvements in sales growth. Instead, we discovered through our T2 tests that sales

growth in not-yet-certified firms was consistently better than non-certified firms, and similar to the results of the T1 test on certified firms. Our tests thus meet the condition for reverse attribution, which indicates that causes other than QCert are responsible for the greater sales growth.

Our findings concerning profitability (ROA) follow a similar pattern. Our earlier cross sectional study (Heras *et al.*, 2002a) indicated that there was an association between profitability and certification. However, on testing for reverse attribution, we found no evidence to support any causal link between ISO 9001 registration and improvements in profitability. Instead we discovered that profitability of the certified firms was consistently better than that of non-certified firms' pre (T1 test) and post (T2 test) registration. Our profitability tests thus meet the condition for reverse attribution, and indicate that better profitability is due to causes other than QCert.

Our findings suggest that cross sectional studies could well be based on a suspect assumption of causality. They ignore the better performance preceding certification that we have found. Our findings on sales growth and profitability clearly illustrate how cross sectional analysis can lead to erroneous inferences of causality, a question we will now discuss in greater detail.

Our re-analysis of our earlier data shows that superior performance precedes QCert rather than the other way around. How does this compare to the other longitudinal studies discussed earlier and summarized in Figure 2? Our findings of T2 test differences indicating the influence of reverse attribution are consistent with the results from Häversjö (2000) and Wayhan *et al.* (2002) shown in Figure 2. Although detailed results for prior performance are not reported by Corbett *et al.* or Naveh and Marcus, both note that the average performance of not-yet certified firms is better than that of non-certified firms (Corbett *et al.*, 2005, p. 1057; Naveh and Marcus, 2005, p. 19). Therefore, there is evidence in all the longitudinal studies of a reverse attribution mechanism being present. However, whereas Häversjö and Wayhan *et al.* show no T3 effect, which supports a dominant influence for a reverse attribution mechanism, Corbett *et al.* and Naveh and Marcus do find a T3 effect which is consistent with the operation of both forward and reverse attribution. In other words, they found improvements in performance were achieved after the registration year when they controlled for better prior performance.

So, in summary, all the longitudinal studies (Häversjö, 2000; Wayhan *et al.*, 2002; Corbett *et al.*, 2005; Naveh and Marcus, 2005) support the existence of the reverse attribution mechanism that we have found in our attribution tests. Our own and two other studies (Häversjö, 2000; Wayhan *et al.*, 2002) suggest that this reverse attribution mechanism was the major influence. Two other studies (Corbett *et al.*, 2005; Naveh and Marcus, 2005) note that this reverse attribution mechanism exists but do not use methods that allow them to judge the size of its influence compared to the QCert effect that they found. In Corbett *et al.* modest performance gains were found while Naveh and Marcus found more substantial gains. However, our own and the other two studies (Häversjö, 2000; Wayhan *et al.*, 2002) suggest that such gains from QCert are far from certain.

Therefore, we must conclude from our analyses that all the studies that could test it found that reverse attribution is a substantial influence in QCert performance attribution. This suggests that where performance benefits are associated with QCert possibly only a modest proportion of the effect found can be reasonably attributed to the causal influence of QCert.

So, what are the possible explanations that might underpin this reverse attribution mechanism? One possibility is that the quality system implementation process takes place well before the certification date and so benefits accrue before registration is achieved. If this is true then event studies should show a discernable step-change in performance leading up to the registration year. Examination of the event study data in Figure 2 for Häversjö and Wayhan *et al.* indicates that there is no evidence to support this explanation. Although year-on-year results vary there is no indication of sustained improvement prior to the certification year. Similarly, when the data analysed here is presented as an event study (Heras *et al.*, 2002b) no support is found for this explanation. In contrast there is support for this explanation in the panel studies of Naveh and Marcus and Corbett *et al.* that do report benefits in the year(s) prior to certification. So, on balance, this explanation for reverse attribution is possible but incomplete.

A simple explanation could be that the cost of accreditation is easier to bear for more profitable firms since they will find the costs easier to absorb making them more inclined to pursue QCert than less profitable firms.

An alternative interpretation is that since all the studies are in the earlier years of ISO 9001 growth in adoption, it could be that these pioneer certified companies are characterized by having a greater exposure to international trade. Thus, these firms are exposed to international standards of competition, and to compete they may already have in place many of the characteristics of “best practice” systems of quality management, prior to seeking accreditation. Therefore, pre and post certification business performance will not differ much, since gaining the “badge of quality” is only recognising what were already good quality management systems.

Or could it be that there are latent common causes to QCert and better performance? An explanation could be that when firms already have in place good quality systems they are more likely to pursue certification early since their costs of implementation are lower and these extant quality systems lead to their better than average performance. There appears to be some support for this explanation since there is generally agreement in the literature on quality management system characteristics (the most dominant being improved conformance quality) that reduce internal costs, or are associated with business performance improvement (Maani *et al.*, 1994; Flynn *et al.*, 1995, 1997; Forker *et al.*, 1996; Caruana and Pitt, 1997; Adam *et al.*, 1997; Samson and Terziovski, 1999; Hendricks and Singhal, 2001; Kaynak, 2003; York and Miree, 2004). This explanation appears to be supported, but once again, the research just quoted can only indicate association, so it may be that this performance precedes the cause as York and Miree (2004) have suggested.

Taking the causal chain to its logical end point suggests that high-performing firms may have a propensity to continually seek and learn from progressive practices/systems (as suggested by Easton and Jarrell, 1998) that can improve and sustain their capabilities. This may explain their above average performance (on the role of learning see Naveh *et al.*, 2004). These characteristics equate to those of Hayes and Wheelwright (1984) Stage Four companies where operations are creative and proactive in developing and adopting new practices and systems that relate to competitive performance (Flynn *et al.*, 1999).

So what are the implications of the reverse attribution mechanism for the interpretation of the literature that is based on forward causation? Clearly, the analyses

cast doubt on any simple inferences about causality being drawn from the broad literature that finds an association of ISO 9001 accreditation with better business performance, since it indicates that firms' superior performance can precede the pursuit of QCert thus inflating any performance difference observed post-certification. Earlier, our review of the empirical literature (Table I) found evidence from empirical research to suggest benefits from QCert were uncertain unless firms had internal or developmental motives to pursue QCert. Could this also be a false attribution of cause, since the possibility exists that these internal/developmental characteristics are those of already high-performing firms and may well contribute to that high performance? Thus, the "internal or developmental motives" intermediate variables may be the common cause of better than average performance that precedes QCert. This is then the reason it is found in the literature as a variable correlated with QCert and better business performance.

Conclusions

In this research, we have put forward three causality attribution tests (Figure 3) that we suggest are necessary if correct attribution of causation is to be made. We have used these tests to analyze our earlier research data on sales growth and profitability of 800 firms which were divided into three samples: certified, not-yet-certified and non-certified over a period of five years. We have shown that the substantial difference between certified and non-certified firms' sales and profitability that we reported in our earlier research (Heras *et al.*, 2002a) cannot be attributed to QCert. The findings of our attribution tests show that better performance preceded QCert indicating the existence of a reverse attribution mechanism.

These findings of better performance preceding registration have also been indicated in the data of all the empirical studies that could be tested for causal attribution (Häversjö, 2000; Wayhan *et al.*, 2002; Corbett *et al.*, 2005; Naveh and Marcus, 2005). We do not exclude the possibility that benefits can be gained from QCert since two (Corbett *et al.*, 2005; Naveh and Marcus, 2005) of the five studies that can indicate causal attribution show that gains were found. However, the lack of such gains in the other three studies does suggest that such gains from QCert are uncertain.

In the discussion, we put forward three possible reasons for the superior performance we have found prior to accreditation to ISO 9001. It could be that the systems required by ISO certification are costly to implement and maintain, so more profitable firms, are more likely to be able to afford it. Alternatively it may be that the certified companies are characterized by having a greater exposure to international trade, and to compete they may have already emulated "best practice" systems of quality management prior to seeking accreditation. A broader explanation is that there is a latent common cause. The common cause could be that high-performing firms are more likely to seek new practices/systems that can improve their capabilities, which ultimately create their above average performance. Thus, better performance is not caused by any single system or practice but is the cumulative result of a process of continuous adoption, learning and adaptation of new management practices/systems.

We are not suggesting that the attribution tests that we have proposed are the ultimate solution in attribution testing, for that advanced computational methods, very large samples and certain conditions must be met (Pearl, 2005). Rather they represent the minimum needed to detect the potential influence of reverse attribution that is due to latent common causes. We concede that we are basing our conclusions on the results

of studies that have used different methods for selecting their control groups. Indeed, Corbett *et al.*'s and Wayhan *et al.*'s results show that different control group selection criteria do influence results. However, regardless of the control group criteria used, no great differences that would alter their overall results were found, so we do not believe that differences in control group methods have any substantial influence that limits our overall findings. Although we base our argument on only five longitudinal studies, these all used actual financial results which provide more reliable evidence than self-reported results. All have shown better performing companies self-select to adopt quality certification. This is true in three very different countries, Denmark, Spain and the USA, which enables us to conclude that the existence of a reverse attribution mechanism is not just a national phenomenon. However, given that over 160 countries with varied cultural and economic regimes have firms registered to the ISO 9001 standard we accept that this self-selection behaviour may not be universal.

So, in summary, the attribution testing method proposed avoids some of the demands of panel matching methods but does introduce the drawback associated with any research that uses cross-sectional control groups. However, the attribution methods strength is that it provides the researcher with a metric for measuring the effect size of reverse attribution.

For researchers the paper provides a method for testing the influence of reverse attribution. It also demonstrates the potential confusion in the attribution of causation in research designs that assume forward causation. The influence of reverse attribution, has we believe, profound implications for the interpretation of causation in the substantial literature that shows QCert is associated with improved business performance. In these studies when a link between business performance improvement and quality management system adoption is found, it is tempting to infer that performance improvement can be attributed to the quality management system change (combined with some intermediate variable(s)). Clearly, the evidence presented here for the presence of a reverse attribution mechanism suggests that counter-intuitive explanations can be equally valid. This suggests that co-causation or reverse attribution deserve wider consideration in the development of explanatory models of performance improvement. We suggest that by adopting research designs that can explicitly measure both causal directions a broader understanding of reverse attribution influences can be established. This could provide the justification for future research into exploration of possible underlying causes. This in turn could lead to the development of broader theory that will enrich our understanding of the complexity of performance attribution.

For practitioners, our findings should give pause for thought. It is indeed tempting for managers to believe that ISO 9001 certification will lead to business benefits. After all, firms that they would like to emulate in terms of performance often have it! This is then reinforced by the seemingly pervasive belief (often incorrectly quoted by certifying bodies as supported by research) that a certified quality management system will increase sales and improve profitability. However, our findings indicate that it might be a wise decision to only pursue accreditation if major customers mandate that it is required, since we have found no conclusive evidence that sales or profitability improve after certification. Most reports indicate that certification is a major investment (Casadesús and Karapetrovic, 2005), yet our findings show that the money spent on certification has not adversely affected the profitability of firms. This does

suggest that cost benefits arising from certification are on average sufficient to offset the investment. Therefore, we are not suggesting to practitioners that certification is a bad investment, rather that inflated expectations of performance improvement are likely to be unfounded.

To summarize, we have explained a method that can be used to indicate what proportion of performance can be attributed to QCert compared to reverse attribution influences. Using this method, we have found evidence from the results of our own and four other longitudinal studies that superior performance was present prior to accreditation in all of the studies. In only two studies of the five were any additional sales or profitability found that could be attributed to QCert. Thus, our findings cast doubt on any simple inference about the direction of causation from the broad literature that finds an association of ISO 9001 accreditation with better business performance. Overall, we have found less evidence for the influence of QCert on business performance, than for the counter-intuitive effect of reverse attribution; a mechanism whose underlying causes need to be investigated in future research.

References

- Abraham, M., Crawford, J., Carter, D. and Mazotta, F. (2000), "Management decisions for effective ISO 9000 accreditation", *Management Decision*, Vol. 38 No. 3, pp. 182-93.
- Adam, E.E. Jr, Corbett, L.M., Flores, B.E., Harrison, N.J., Lee, T.S., Rho, B., Ribera, J., Samson, D. and Westbrook, R. (1997), "An international study of quality improvement approach and firm performance", *International Journal of Operations & Production Management*, Vol. 17 No. 9, pp. 842-73.
- Ahire, S.L., Landeros, R. and Golhar, D.Y. (1995), "Total quality management: a literature review and an agenda for future research", *Production and Operations Management*, Vol. 4 No. 3, pp. 277-300.
- ANAB (2005), "What are the benefits of certification?", available at: www.anab.org/HTMLfiles/ab_faq.htm (accessed December 2005).
- Barber, B.M. and Lyon, J.D. (1996), "Detecting abnormal performance: the empirical power and specification of test statistics", *Journal of Financial Economics*, Vol. 41, pp. 359-99.
- Blaikie, N. (2003), *Analyzing Quantitative Data*, Sage, London.
- Breeze, S. (2004), "Raise the standard to win worldwide success", *Daily Telegraph*, 20 May.
- Briscoe, J.A., Fawcett, S.E. and Todd, R.H. (2005), "The implementation and impact of ISO 9000 among small manufacturing enterprises", *Journal of Small Business Management*, Vol. 43 No. 3, pp. 309-30.
- Brown, A., van de Wiele, T. and Loughton, K. (1998), "Smaller enterprises' experience with ISO 9000", *International Journal of Quality & Reliability Management*, Vol. 15 No. 3, pp. 273-85.
- Buttle, F. (1997), "ISO 9000: marketing motivations and benefits", *International Journal of Quality & Reliability Management*, Vol. 14 No. 9, pp. 939-47.
- Carson, P.P., Lanier, P.A., Carson, K.D. and Guidry, B.N. (2000), "Clearing a path through the management fashion jungle: some preliminary trailblazing", *Academy of Management Journal*, Vol. 43 No. 6, pp. 1143-56.
- Caruana, A. and Pitt, L. (1997), "INTQUAL – an internal measure for service quality and the link between service quality and business performance", *European Journal of Marketing*, Vol. 31 No. 8, pp. 604-17.

- Casadesús, M. and Karapetrovic, S. (2005), "Has ISO 9000 lost some of its luster? A longitudinal impact study", *International Journal of Quality & Reliability Management*, Vol. 25 No. 6, pp. 508-96.
- Chou-Chua, C., Goh, M. and Wan, T.B. (2003), "Does ISO 9000 certification improve business performance?", *International Journal of Quality & Reliability Management*, Vol. 20 No. 8, pp. 936-53.
- Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences*, Lawrence Erlbaum Associates, Hillsdale, NJ.
- Corbett, C.J., Montes-Sancho, M.J. and Kirsch, D.A. (2005), "The financial impact of ISO 9000 certification in the United States: an empirical analysis", *Management Science*, Vol. 51 No. 7, pp. 1046-59.
- Deming, W.E. (1986), *Out of the Crisis*, Center for Advanced Engineering Study, Cambridge, MA.
- Dick, G.P.M., Gallimore, K. and Brown, J.C. (2000), "Does ISO 9000 give a quality emphasis advantage? A comparison of large service and manufacturing organizations", *Quality Management Journal*, Vol. 8 No. 1, pp. 52-61.
- Dimara, E., Tsekouras, K., Skura, D. and Goutsos, S. (2004), "Strategic orientation and financial performance of firms implementing ISO 9000", *International Journal of Quality & Reliability Management*, Vol. 21 No. 1, pp. 72-89.
- Easton, G.S. and Jarrell, S.L. (1998), "The effects of total quality management on corporate performance: an empirical investigation", *Journal of Business*, Vol. 71 No. 2, pp. 253-65.
- Ebrahimpour, M., Withers, B.E. and Hikmet, N. (1997), "Experiences of US and foreign-owned firms: a new perspective on ISO 9000 implementation", *International Journal of Production Research*, Vol. 35 No. 2, pp. 569-76.
- Flynn, B.B., Schroeder, R.G. and Flynn, E.J. (1999), "World class manufacturing: an investigation of Hayes and Wheelwright's foundation", *Journal of Operations Management*, Vol. 17, pp. 249-69.
- Flynn, B.B., Schroeder, R.G. and Sakakibara, S. (1995), "The impact of quality management practices on performance and competitive advantage", *Decision Sciences*, Vol. 26 No. 5, pp. 659-92.
- Flynn, B.B., Schroeder, R.G., Flynn, E.J., Sakakibara, S. and Bates, K.A. (1997), "World-class manufacturing project: overview and selected results", *International Journal of Operations & Production Management*, Vol. 17 No. 7, pp. 671-85.
- Forker, L.B., Vickery, S.K. and Droge, C.L. (1996), "The contribution of quality to business performance", *International Journal of Operations & Production Management*, Vol. 16 No. 8, pp. 44-62.
- Garvin, D. (1984), "What does 'product quality' really mean?", *Sloan Management Review*, Fall, pp. 25-43.
- Guenther, D.A. and Rosman, J. (1994), "Differences between Compustat and CRSP SIC codes and related effects on research", *Journal of Accounting Economics*, Vol. 18, pp. 115-8.
- Häversjö, T. (2000), "The financial effects of ISO 9000 registration for Danish companies", *Managerial Auditing Journal*, Vol. 15 Nos 1/2, pp. 47-52.
- Hayes, R.H. and Wheelwright, S.C. (1984), *Restoring our Competitive Edge*, Wiley, New York, NY.
- Hendricks, K.B. and Singhal, V.R. (2001), "Firm characteristics, total quality management, and financial performance", *Journal of Operations Management*, Vol. 19, pp. 239-85.
- Heras, I., Casadesús, M. and Dick, G.P.M. (2002a), "ISO 9000 certification and the bottom line: a comparative study of the profitability of Basque region firms", *Managerial Auditing Journal*, Vol. 17 Nos 1/2, pp. 72-8.

-
- Heras, I., Dick, G.P.M. and Casadesús, M. (2002b), "ISO 9000 registration's impact on sales and profitability: a longitudinal analysis of performance before and after accreditation", *International Journal of Quality & Reliability Management*, Vol. 19 No. 6, pp. 774-91.
- Huang, F., Hong, C. and Chen, C. (1999), "A study of ISO 9000 process, motivation and performance", *Total Quality Management*, Vol. 10 No. 7, pp. 1009-25.
- ISO (2005), *The ISO Survey of ISO 9001:2000 and ISO 14001 Certificates (14th Cycle)*, International Organizations for Standardisation, Geneva.
- Jones, R., Arndt, G. and Kustin, R. (1997), "ISO 9000 among Australian companies: impact of time and reasons for seeking certification on perceptions of benefits received", *International Journal of Quality & Reliability Management*, Vol. 14 No. 7, pp. 650-60.
- Kaynak, H. (2003), "The relationship between total quality management practices and their effects on firm performance", *Journal of Operations Management*, Vol. 21, pp. 405-25.
- Lima, M.A.M., Resende, M. and Hasenclever, L. (2000), "Quality certification and performance of Brazilian firms: an empirical study", *International Journal of Production Economics*, Vol. 66, pp. 143-7.
- Maani, K.E., Putterill, M.S. and Sluti, D.G. (1994), "Empirical analysis of quality improvement in manufacturing", *International Journal of Quality & Reliability Management*, Vol. 11 No. 7, pp. 19-37.
- Naveh, E. and Marcus, A. (2005), "Achieving competitive advantage through implementing a replicable management standard: installing and using ISO 9000", *Journal of Operations Management*, Vol. 24 No. 1, pp. 1-26.
- Naveh, E., Marcus, A. and Moon, H.K. (2004), "Implementing ISO 9000: performance improvement by first and second movers", *International Journal of Production Research*, Vol. 42 No. 9, pp. 1846-63.
- Pearl, J. (2005), *Causality: Models Reasoning and Inference*, Cambridge University Press, Cambridge, MA.
- Prabhu, V., Appleby, A., Yarrow, D. and Mitchell, E. (2000), "The impact of ISO 9000 and TQM on best practice/performance", *TQM Magazine*, Vol. 12 No. 2, p. 8491.
- Samson, D. and Terziovski, M. (1999), "The relationship between quality management practices and operational performance", *Journal of Operations Management*, Vol. 17, pp. 393-409.
- Santos, L. and Escanciano, C. (2001), "Benefits of the ISO 9000: 1994 system", *International Journal of Quality & Reliability Management*, Vol. 19 No. 3, pp. 321-44.
- Simmons, D.A. (1999), "Health care quality and ISO 9000", paper presented at University Health Care Consortium Annual Meeting, April 27.
- Singles, J., Ruel, G. and van de Water, H. (2001), "ISO 9000 series – certification and performance", *International Journal of Quality & Reliability Management*, Vol. 18 No. 1, pp. 62-75.
- Sousa, R. and Voss, C.A. (2002), "Quality management re-visited: a reflective review and agenda for future research", *Journal of Operations Management*, Vol. 20, pp. 91-109.
- Sun, H. (2000), "The patterns of implementing TQM versus ISO 9000 at the beginning of the 1990s", *International Journal of Quality & Reliability Management*, Vol. 16 No. 3, pp. 201-15.
- Terziovski, M., Power, D. and Sohal, A. (2003), "The longitudinal effects of the ISO 9000 certification process on business performance", *European Journal of Operational Research*, Vol. 146, pp. 508-95.
- Terziovski, M., Samson, D. and Dow, D. (1997), "The business value of quality management systems certification: evidence from Australia and New Zealand", *Journal of Operations Management*, Vol. 15 No. 1, pp. 1-18.

- Tsekouras, K., Dimara, E. and Skura, D. (2002), "Adoption of a quality assurance system and its effect on firm performance: a study of Greek firms implementing ISO 9000", *Total Quality Management*, Vol. 13 No. 6, pp. 827-41.
- Tzelepis, D., Tsekouras, K., Skuras, D. and Dimara, E. (2006), "The effects of ISO 9001 on firms' productive efficiency", *International Journal of Operations & Production Management*, Vol. 26 No. 10, pp. 1146-65.
- Uzumeri, M.V. (1997), "ISO 9000 and other management metastandards: principles for management practice", *Academy of Management Executive*, Vol. 11 No. 1, pp. 21-36.
- Wayhan, V., Kirche, E. and Khumawala, B. (2002), "ISO 9000 certification: financial performance implications", *Total Quality Management*, Vol. 13 No. 2, pp. 217-31.
- White, P.A. (1990), "Ideas about causation in philosophy and psychology", *Psychology Bulletin*, Vol. 108 No. 1, pp. 3-18.
- Yahya, S. and Goh, W. (2001), "The implementation of an ISO 9000 quality system", *International Journal of Quality & Reliability Management*, Vol. 18 No. 9, pp. 941-66.
- Yeung, A.C.L., Lee, T.S. and Chan, L.Y. (2003), "Senior management perspectives and ISO 9000 effectiveness: an empirical research", *International Journal of Production Research*, Vol. 41 No. 3, pp. 545-69.
- York, K.M. and Miree, C.E. (2004), "Causation or covariation: and empirical examination of the link between TQM and financial performance", *Journal of Operations Management*, Vol. 22 No. 3, pp. 291-311.